APPENDIX 2

APPENDIX 2

A 0.9-V, 150-MHz, 10-mW, 4 mm^2, 2-D Discrete Cosine Transform Core Processor with Variable Threshold-Voltage (VT) Scheme

Tadahiro Kuroda, *Member, IEEE*, Tetsuya Fujita, Shinji Mita, Tetsu Nagamatsu, Shinichi Yoshioka, Kojiro Suzuki, Fumihiko Sano, Masayuki Norishima, Masayuki Murota, Makoto Kako, Masaaki Kinugawa, *Member, IEEE*, Masakazu Kakumu, *Member, IEEE*, and Takayasu Sakurai, *Member, IEEE*

Abstract—A 4 mm^2, two-dimensional (2-D) 8 × 8 discrete cosine transform (DCT) core processor for HDTV-resolution video compression/decompression in a 0.3-μm CMOS triple-well, double-metal technology operates at 150 MHz from a 0.9-V power supply and consumes 10 mW, only 2% power dissipation of a previous 3.3-V design. Circuit techniques for dynamically varying threshold voltage (VT scheme) are introduced to reduce active power dissipation with negligible overhead in speed, standby power dissipation, and chip area. A way to explore $V_{DD} - V_{th}$ design space is also studied.

I. Introduction

LOWERING both the supply voltage V_{DD} and threshold voltage V_{th} enables high-speed, low-power operation [1], [2]. This approach, however, raises two problems [3], [4], 1) degradation of worst-case speed due to V_{th} fluctuation in low V_{DD}, and 2) increase in standby power dissipation in low V_{th}. To solve these problems, several schemes are proposed. A self-adjusting threshold voltage (SAT) scheme [5] reduces V_{th} fluctuation in an active mode by adjusting substrate bias with a feedback control circuit. A standby power reduction (SPR) scheme [6] raises V_{th} in a standby mode by switching substrate bias between the power supply and an external additional supply higher than V_{DD} or lower than GND. A multi threshold voltage CMOS (MT-CMOS) scheme [7] employ low V_{th} for fast circuit operation and high V_{th} for providing and cutting internal supply voltage. The SAT and the SPR are both based upon the same idea that V_{th} is controlled dynamically through substrate bias. However, the two schemes cannot be combined because the SPR requires the external supply for the substrate bias while the SAT generates the substrate bias internally. The MT-CMOS does not solve the first problem. It requires very large transistors for the internal power supply control to impose area and yield penalties, otherwise degrading circuit speed. Furthermore, it cannot be applied to memory elements without circuit tricks which add another area and speed penalties.

This paper presents a variable threshold voltage scheme (VT scheme) which can solve these two problems uniformly in a unified way by controlling substrate bias with substrate bias feedback control circuits. Unlike the conventional approaches, it requires no external power supply for the substrate bias, leaves no restriction in use, imposes practically no penalty in speed and chip area, and can be applied to both logic gates and memory elements. The VT scheme is employed in a two-dimensional (2-D) 8 × 8 discrete cosine transform (DCT) core processor for portable HDTV-resolution video compression/decompression. This DCT in a 0.3-μm CMOS technology operates at 150 MHz from a 0.9-V power supply and consumes 10 mW, only 2% power dissipation of a previous 3.3-V design [8].

In Section II, low V_{DD}, low V_{th} design space is explored to investigate V_{th} target. In Section III, the VT scheme is presented, followed by descriptions of circuit implementations in Section IV. Section V details the design of the DCT. Experimental results appear in Section VI. Section VII is dedicated for conclusions.

II. Exploring Low-V_{DD} Low-V_{th} Design Space

CMOS power dissipation is given by

$$P = \frac{1}{2} \cdot p_t \cdot f_{CLK} \cdot C_L \cdot V_{DD}^2 + I_0 \cdot 10^{-(V_{th}/S)} \cdot V_{DD} \quad (1)$$

where p_t is the switching probability, f_{CLK} is the clock frequency, C_L is the load capacitance, S is the subthreshold swing, and I_0 is a constant which is proportional to total transistor width in a chip. The first term represents dynamic power dissipation due to charging and discharging of the load capacitance, and the second term is leakage current dissipation due to subthreshold conduction. Since the dominant term in a typical CMOS design is the dynamic power dissipation, lowering V_{DD} is effective to low-power design.

Gate propagation delay, on the other hand, is approximately given in [9] by

$$t_{pd} = \frac{k \cdot C_L \cdot V_{DD}}{(V_{DD} - V_{th})^\alpha} \quad (2)$$

Manuscript received April 11, 1996; revised July 23, 1996.

T. Kuroda, T. Fujita, S. Mita, T. Nagamatsu, S. Yoshioka, K. Suzuki, F. Sano, and T. Sakurai are with System ULSI Engineering Laboratory, Toshiba Corp., Kawasaki, Japan.

M. Norishima and M. Kakumu are with LSI Div. II, Toshiba Corp., Kawasaki, Japan.

M. Murota, M. Kako, and M. Kinugawa are with ULSI Device Engineering Laboratory, Toshiba Corp., Kawasaki, Japan.

Publisher Item Identifier S 0018-9200(96)07943-7.

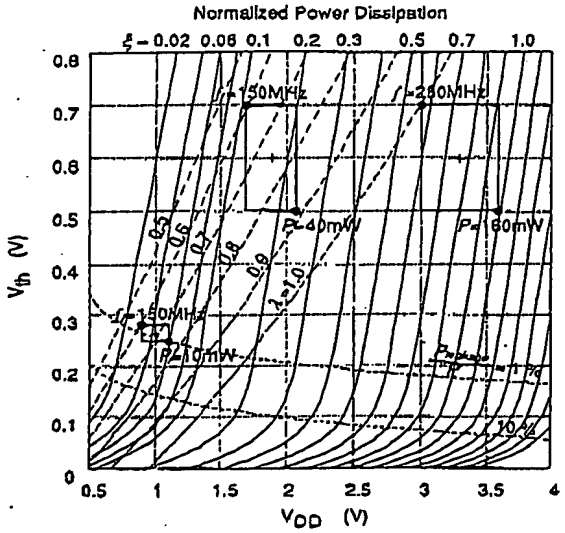

Fig. 1. Exploring low-V_{DD}, low-V_{th} design space. Contour lines in terms of speed (broken lines) and power (solid lines) are drawn.

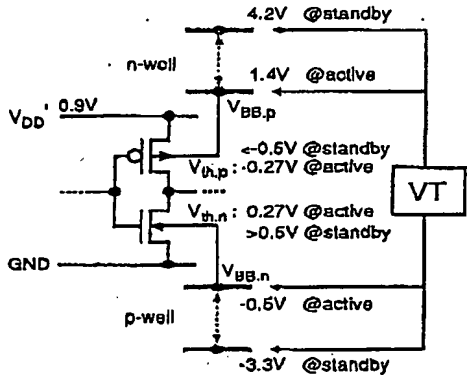

Fig. 2. Variable threshold-voltage (VT) scheme.

100 μA from the substrate to lower V_{BB} using a 50 MHz ring oscillator. This current is large enough for V_{BB} to settle down within 10 μs after a power-on. When V_{BB} goes lower than $V_{active(+)}$, the pump driving frequency drops to 5 MHz and the SSB draws 10 μA to control V_{BB} more precisely. The SSB stops when V_{BB} drops below V_{active}. V_{BB}, however, rises gradually due to device leakage current through MOS transistors and junctions, and reaches V_{active} to activate the SSB again. In this way, V_{BB} is controlled at V_{active} by the on–off control of the SSB. When V_{BB} goes deeper than $V_{active(-)}$, the SCI turns on to inject 30 mA into the substrate. Therefore, even if V_{BB} jumps beyond $V_{active(+)}$ or $V_{active(-)}$ due to a power line bump for example, V_{BB} is quickly recovered to V_{active} by the SSB and the SCI. When "SLEEP" signal is asserted ("1") to go to the standby mode, the SCI is disabled and the SSB is activated again and 100 μA current is

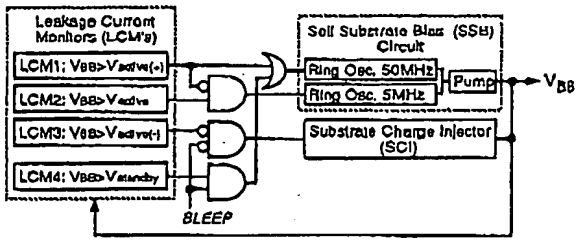

Fig. 3. VT block diagram.

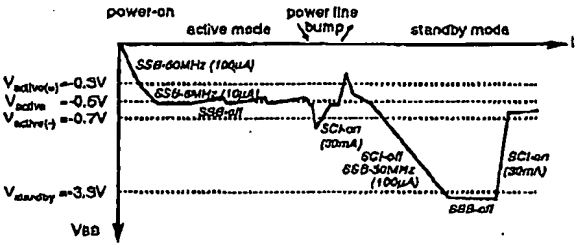

Fig. 4. Substrate-bias control in VT.

drawn from the substrate until V_{BB} reaches $V_{standby}$. V_{BB} is controlled at $V_{standby}$ in the same way by the on–off control of the SSB. When "SLEEP" signal becomes "0" to go back to the active mode, the SSB is disabled and the SCI is activated. The SCI injects 30 mA current into the substrate until V_{BB} reaches $V_{active(-)}$. V_{BB} is finally set at V_{active}. In this way, the SSB is mainly used for a transition from the active mode to the standby mode, while the SCI is used for a transition from the standby to the active mode. An active to standby mode transition takes about 100 μs, while a standby to active mode transition is completed in 0.1 μs. This "slow falling asleep but fast awakening" feature is acceptable for most of the applications.

The SSB operates intermittently to compensate for the voltage fluctuation in the substrate due to the substrate current in the active and the standby modes. It therefore consumes several microamperes in the active mode and less than one nanoampere in the standby mode, both much lower than the chip power dissipation. Energy required to charge and discharge the substrate for switching between the active and the standby modes is less than 10 nJ. Even when the mode is switched 1000 times in a second, the power dissipation becomes only 10 μW. The leakage current monitor should be designed to dissipate less than 1 nA because it always works even in the standby mode. The low-power circuit design technique is described in the next section.

IV. Circuit Implementations

A. Leakage Current Monitor (LCM)

The substrate bias is generated by the SSB which is controlled by the leakage current monitor (LCM). The LCM is therefore a key to the accurate control in the VT scheme. Fig. 5

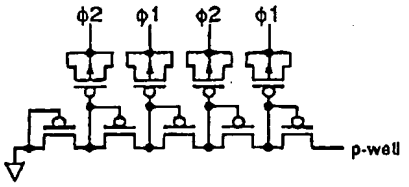

Fig. 7. Pump circuit in SSB.

Compared with the conventional LCM where V_b is generated by dividing the V_{DD}-GND voltage with high impedance resistors, the V_{th} controllability including the static and dynamic effects is improved from ±0.05 V to less than ±0.01 V, response delay is shortened from 0.6 to 0.1 μs, and Si area is reduced from 33 250 to 670 μm^2. This layout area reduction is brought by the elimination of the high impedance resistors by polysilicon.

B. Self-Substrate Bias Circuit (SSB)

Fig. 7 depicts a schematic diagram of a pump circuit in the SSB. PMOS transistors of the diode configuration are connected in series whose intermediate nodes are driven by two signals, Φ1 and Φ2, in 180° phase shift. Every other transistor, therefore, sends current alternately from p-well to GND, resulting in lower p-well bias than GND. The SSB can pump as low as −4.5 V. SSB circuits are widely used in DRAM's and E^2PROM's, but two orders of magnitude smaller circuit can be used in the VT scheme. The driving current of the SSB is 100 μA, while it is usually several milliamperes in DRAM's. This is because substrate current generation due to the impact ionization is a strong function of the supply voltage. Substrate current in a 0.9-V DCT is considerably smaller than that in a 3.3-V design. Substrate current introduced from I/O pads does not affect the DCT macro because it is separated from peripheral circuits by a triple-well structure. Eventually, no substrate current is generated in the standby mode. From these reasons, the pumping current in the SSB can be as small as several percent of that in DRAM's. Silicon area is also reduced considerably. Another concern about the SSB is an initialization time after a power-on. Even in a 10 mm square chip, V_{BB} settles down within 200 μs after a power-on, which is acceptable in real use.

C. Substrate Charge Injector (SCI)

In the VT scheme, care should be taken so that no transistor sees high-voltage stress of gate oxide and junctions. Transistors are optimized for use at 3.3 V. The gate oxide thickness is 8 nm. The maximum voltage that assures sufficient reliability of the gate oxide is $V_{DD}+20\%$, or 4 V. The SCI in Fig. 8 receives a control signal that swings between V_{DD} and GND at node N_1 to drive substrate from $V_{standby}$ to V_{active}. In the standby-to-active transition, $V_{DD}+|V_{standby}|$ that is about 6.6 V at maximum can be applied between N_1 and N_2. However, as shown in SPICE simulated waveforms in Fig. 8, $|V_{GS}|$ and $|V_{GD}|$ of M1 and M2 never exceeds the larger of V_{DD} and $|V_{standby}|$. All other transistors in the VT circuit and

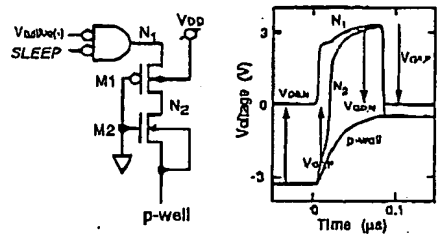

Fig. 8. SCI and its waveforms simulated by SPICE.

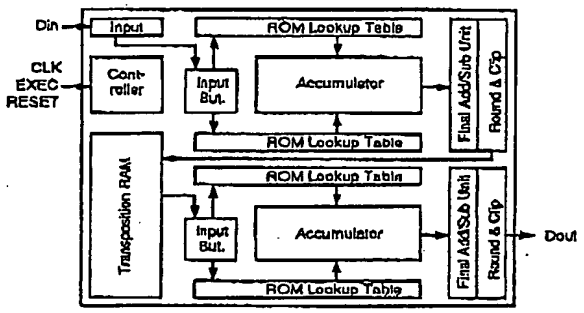

Fig. 9. DCT block diagram.

the DCT macro receive $(V_{DD}-V_{th})$ on their gate oxide when the channel is formed in the depletion and the inversion mode, and less than $|V_{standby}|$ in the accumulation mode. These considerations lead to a general guideline that $V_{standby}$ should be limited to $-(V_{DD}+20\%)$. $V_{standby}$ of $-(V_{DD}+20\%)$, however, can shift V_{th} big enough to reduce the leakage current in the standby mode. The body effect coefficient, γ, can be adjusted independently to V_{th} by controlling the doping concentration density in the channel-substrate depletion layer.

V. DCT Design

A. Circuit Design

This DCT core processor executes 2-D 8 × 8 DCT and inverse DCT. A block diagram is illustrated in Fig. 9. The DCT is composed of two one-dimensional (1-D) DCT and inverse DCT processing units and a transposition RAM. Rounding circuits and clipping circuits which prevent overflow and underflow are also implemented in the cell. The DCT has a concurrent architecture based on distributed arithmetic and a fast DCT algorithm, which enables high throughput DCT processing of one pixel per clock. It also has fully pipelined structure. The 64 input data sampled in every clock cycles are outputted after 112 clock cycle latency.

Various memories which use the same low V_{th} transistors as logic gates are employed in the DCT. Table lookup ROM's (16 b × 32 words × 16 banks) employ contact programming and an inverter-type sense-amplifier. Single-port SRAM's (16 b × 64 words × 2 banks) and dual-port SRAM's (16 b × 8 words × 2 banks) employ a six-transistor cell and a latch sense-amplifier. They all exhibit wide operational margin in

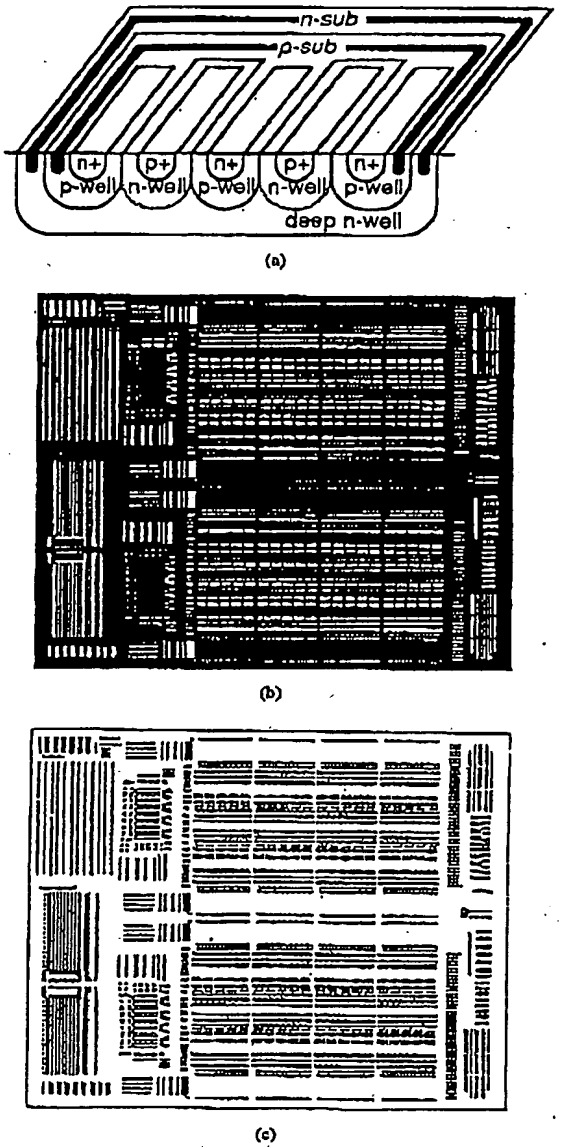

Fig. 11. DCT layout modification for the VT scheme: (a) device cross-section, (b) p-well (one island), and (c) n-well (pieces of islands) in deep n-well.

TABLE I
FEATURES

Technology	0.3 μm CMOS, triple-well, double-metal, T_{ox} = 8 nm, V_{th} = 0.15 V ± 0.1 V
Power supply voltage	1.0 V ± 0.1 V
Power dissipation	10 mW @ 150 MHz
Standby current	<10 nA @ 70°C
Transistor count	120K Tr
Area	2.0 × 2.0 mm^2
Function	8 × 8 DCT and inverse DCT
Data format	9-b signed (pixel), 12-b signed (DCT)
Latency	112 clocks
Throughput	64 clocks/block
Accuracy	CCITT H.261 compatible

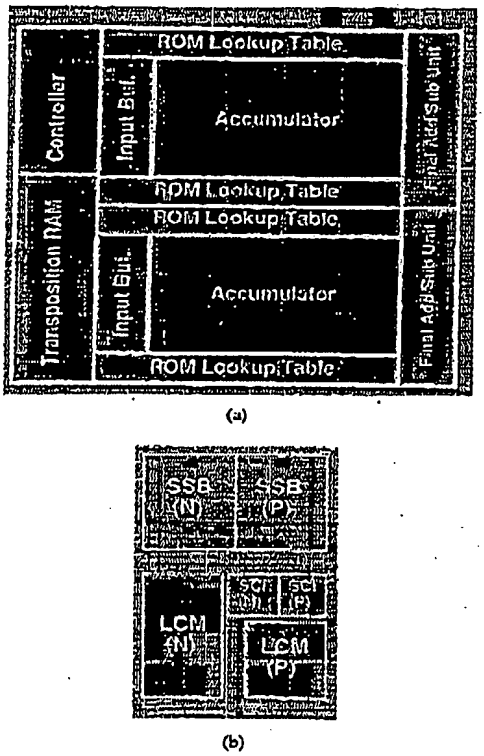

Fig. 12. Chip micrograph; (a) DCT macro and (b) VT circuits.

even when 100 kΩ resistance is added between the substrate and the output of the SSB.

VII. Conclusions

A 4 mm^2 2-D DCT core processor for portable multimedia equipment with HDTV-resolution video compression and decompression has been developed in a 0.3-μm CMOS, triple-well, double-metal technology. It operates at 150 MHz from a 0.9 V power supply and dissipates 10 mW, which is only 2% of the previous 3.3 V design. Circuit design techniques for dynamically varying threshold voltage (VT scheme) are introduced to reduce active power dissipation with negligible overhead in speed, standby power dissipation, and chip area. The active-to-standby mode transition takes 120 μs, while the standby-to-active mode transition is completed within 0.2 μs. The VT scheme can be applied to both logic gates and memory elements. Generation of the low-voltage V_{DD}' on chip is a future research work.

Shinji Mita was born in Aichi, Japan, on March 18, 1970. He received the B.S. degree in electrical engineering from the University of Kyushu, Fukuoka, Japan, in 1992.

In 1992, he joined Toshiba Corporation, Kawasaki, Japan. Since 1992, he has been with Semiconductor Device Engineering Laboratory at Toshiba, where he has been involved in the research and development of multimedia LSI's. His current interests include high-speed, low-power, low-voltage techniques in CMOS.

Tetsu Nagamatsu was born in Yamaguchi, Japan, on August 13, 1960. He received the B.S. degree in applied physics from Waseda University in 1984 and the M.S. degree in energy science from Tokyo Institute of Technology.

He joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kanagawa, Japan, in 1986. He was engaged in the research and development of BiCMOS logic gate, GA and memory macros. Then he was also engaged in the design of DCT/IDCT macro for MPEG2 Decoder. He currently belongs to the System LSI Laboratory. And he is engaged in the development of high-speed CMOS differential circuits for telecommunication area.

Shinichi Yoshioka was born in Tokyo, Japan, in 1963. He received the B.E. and M.E. degrees in electric engineering from Keio University, Tokyo, Japan, in 1987 and 1989, respectively.

In 1989, he joined the Research and Development Center, Toshiba Corporation, Kawasaki, Japan, where he has been working on logic LSI's. Since 1996, he has been engaged in the development of ULSI processors at the System ULSI Laboratory, Toshiba Corporation.

Kojiro Suzuki was born in Kawasaki, Japan, on October 11, 1967. He received the B.S., M.S., and Ph.D. degrees in superconductivity from University of Tokyo, Tokyo, Japan, in 1990, 1992, and 1995, respectively. His Ph.D. work was on design and fabrication of a high-sensitivity SQUID with $Nb/AlO_x/Nb$ Josephson junctions.

In 1995, he joined Toshiba Corporation, Kawasaki, Japan, where he was engaged in the design of low-voltage bus circuits, supply voltage controller, and the research of CMOS low-power design rules. His present interest is low-power, low-voltage techniques in CMOS circuits.

Fumihiko Sano was born in Shiga, Japan, on March 18, 1967. He received the B.S. degree in electrical engineering from Fukui National College of Technology, Japan, in 1988.

He joined Toshiba Microelectronics Corporation, Kawasaki, Japan. He then joined Toshiba's Microelectronics Engineering Laboratory, Kawasaki, Japan, where he has been engaged in the research and development of BiCMOS macrocells for high-performance ASIC's. He has also been engaged in the research and development of VLD macrocells implemented in MPEG2-decoder LSI.

Masayuki Norishima was born in Tokyo, Japan, on January 6, 1962. He received the B.S. degree in pure and applied science from Tokyo University, Tokyo, Japan, in 1986.

He joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, in 1986. From 1986 to 1990, he was engaged in the research and development of process/device technology for Bi-CMOS logic VLSI. From 1990 to 1995 he was engaged in the research and development of process/device technology for high speed CMOS logic VLSI. In September 1995, he joined the LSI Division 2, Toshiba Corporation, Kawasaki, Japan, where he is working on the development of process/device technology for CMOS logic VLSI, focused on mass production in fabs.

Mr. Norishima is a member of the Japan Society of Applied Physics.

Masayuki Murota was born in Kanagawa, Japan. He received B.E. and M.E. degrees in electrical engineering from Hosei University, Tokyo, Japan in 1988 and 1990, respectively.

He joined the Microelectronics Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, in 1990. He had been engaged in the research and development for process technology of multilevel interconnection of LSI's. Since 1995, he has been engaged in the research and development of high performance microprocessor and logic devices.

Mr. Murota is a member of the Japan Society of Applied Physics.

Makoto Kuko was born in Aichi, Japan, in 1972. He graduated from the Technical High School of Higashiyama, Aichi, Japan, in 1991.

He joined the Toshiba Corporation, Kawasaki, Japan, in 1991. From 1991 to 1992, he attended a one-year technical training program at the Toshiba Computer School, Kawasaki, Japan. He has been engaged in the research and development of process technology for high performance microprocessor and logic devices.

Masaaki Kinugawa (M'91) was born in Hyougo, Japan, in 1958. He received the B.S. degree in physics from Kyoto University, Kyoto, Japan, in 1981 and the M.S. degree in applied physics from Tokyo University, Tokyo, Japan in 1983.

In 1983, he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan. He was engaged in the development of CMOS process and device technology for static RAM and high performance RISC chips. He moved to the ULSI Device Engineering Laboratory in 1996.

Mr. Kinugawa is a member of IEEE Electron Device Society.